

# Optimization in telecommunication networks

Citation for published version (APA):

van Hoesel, C. P. M. (2004). *Optimization in telecommunication networks*. METEOR, Maastricht University School of Business and Economics. METEOR Research Memorandum No. 045  
<https://doi.org/10.26481/umamet.2004045>

**Document status and date:**

Published: 01/01/2004

**DOI:**

[10.26481/umamet.2004045](https://doi.org/10.26481/umamet.2004045)

**Document Version:**

Publisher's PDF, also known as Version of record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Optimization in Telecommunication Networks

Stan van Hoesel

November 10, 2004

## Abstract

Network design and network synthesis have been the classical optimization problems in telecommunication for a long time. In the recent past, there have been many technological developments such as digitization of information, optical networks, internet, and wireless networks. These developments have led to a series of new optimization problems. This manuscript gives an overview of the developments in solving both classical and modern telecom optimization problems.

We start with a short historical overview of the technological developments. Then, the classical (still actual) network design and synthesis problems are described with an emphasis on the latest developments on modelling and solving them. Classical results such as Menger's disjoint paths theorem, and Ford-Fulkerson's max-flow-min-cut theorem, but also Gomory-Hu trees and the Okamura-Seymour cut-condition, will be related to the models described. Finally, we describe recent optimization problems such as routing and wavelength assignment, and grooming in optical networks.

## 1 A short history on technological developments

This section presents an overview of the technological breakthroughs that resulted in the current level of telecommunication. We start with the fundamental developments that led to the basic usage of telephone (wired and wireless), the classical history. Next, we discuss recent developments such as digitization, optical technology and wireless networks.

### 1.1 Classical history

Fast communication over large distances already appealed the ancient Greeks. Without much technical knowledge of electricity and light they had to resort to human labor. It is not until the end of the 18th century before we see real telecommunication: instant communication over large distances, with little human effort. Claude Chappe (in France, in 1793) and Abraham Edelcrantz (in Sweden, one year later) invented the optical telegraph. This device could send light signals over large distances by use of a series of instruments based on mirrors. Besides the technical details both inventors also developed a protocol for transmitting, which decided when, how and where information was sent. It took only a few years before a nation-wide network of telegraphs was implemented in both countries. This, in spite of strong opposition

against the telegraph. Edelcrantz, in his "Treatise on Telegraphs" (1796), dealt with the criticism as follows.

*"It often happens, with regard to new inventions, that one part of the general public finds them useless and another part considers them to be impossible. When it becomes clear that the possibility and the usefulness can no longer be denied, most agree that the whole thing was fairly easy to discover and that they knew was significant".*

From that moment a regular stream of innovations, each with its own practical advantages and mathematical problems, emerged. In 1831 Samuel Morse invented the electrical telegraph, which had a major advantage over the optical telegraph in the sense that no human effort was necessary for operating the intermediate mirrors. Moreover, Morse drastically improved the transmission protocol by introducing an efficient and simple encoding of the alphabet. The next innovation was logical but took another 45 years. In 1876, Alexander Graham Bell invented the telephone. This device replaced the binary encoding of telegraphs with human voice, by translating voice into electrical signals and back. Bell himself, in all modesty, described the value of his invention like this:

*"This invention's greatest advantage over every other form of electrical apparatus is the fact that it could be used by anyone, as all other telegraphic machines produce signals which require to be translated by experts, and such instruments are therefore extremely limited in their application".*

Pretty soon telephone networks appeared in several cities in the United States, starting in New Haven in 1879. Bell and the first telephone companies had an interesting controversy, related to the structure of these networks. Bell wanted every customer to be connected with all other customers by a dedicated circuit, whereas the companies opted for a centralized model, where a centrally placed switch was connected with the customers, see Figure 1. The advantages of the second method are clear: the number of connections grows linearly instead of quadratically with the number of customers. The disadvantage is that the number of calls that can be handled simultaneously by the switch is limited (to one in the very beginning). Thus, there is no guaranteed connection when a customer requires one.

After the city networks were established, these networks were interconnected with one another, introducing multiple levels in the network structure. A circuit connecting customers may be built up of several links. Thus, the distance between customers, measured in number of links to be passed by a circuit, is therefore also an important measure. Moreover, when more and more customers start using the network, failures become increasingly more problematic. Thus, protection and restoration methods for circuits must be developed. Connectivity plays an important role here. Complete connectivity (full mesh) is hardly realizable nor a necessity, on any level of a network. On lower levels single connectivity generally suffices. Thus, such networks are generally built from units like stars. On higher levels, double or even higher connectivity is required. Nowadays, such networks use the ring as the basic building block, see Figure 1. The technology and accompanying problems led to a series of problems, called network design and synthesis which are still of very high importance when developing a telecommunications network.

The last fundamental technological invention of importance was done by Marconi, who developed wireless telegraphy by use of electromagnetic waves in 1897. This idea's main virtue

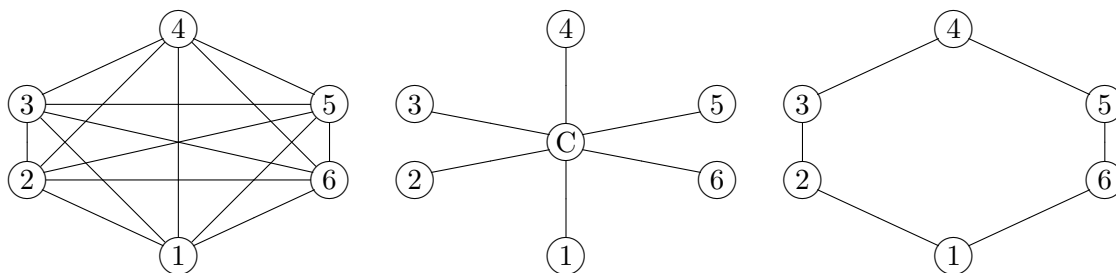


Figure 1: Full mesh, star, ring.

is that for this kind of communication, physical networks are not necessary. This makes it applicable for specific applications, such as ships at sea, planes and military operations. The logical step to speech over wireless connections was made in 1906, when the first radio band wave communication of human speech was accomplished by Reginald Fessenden.

## 1.2 Modern developments

There have been three major technological developments in the past 50 years that greatly influenced the status of current telecommunication. In the 1960s experiments started with digital signal transmission instead of analog transmission. This technology made it practical to transmit not only voice but any kind of digitized information, and thus paved the way for the start of Internet. In the 1980s wireless telephony became practical for consumers, due to miniaturization of mobile telephone components and new compression techniques for digital speech. Finally, the development in the 1990s of optical transmission components to replace electrical equipment allowed for a huge increase in availability of cheap bandwidth. These three developments will be discussed shortly, below.

### 1.2.1 From analog to digital networks

Digitization of speech is done by *sampling*. An analog signal generated by human voice is monitored 8000 times per second. Each of the samples is translated into an 8-bit word of information. Thus, for a single call, a circuit with a bandwidth of 64kbps is required.

Beginning in the 1960s, telephone systems gradually began converting their connections to a digital switching system. The first network protocol for digital information transmission was ISDN (Integrated Services Digital Network), which was standardized in 1984 by the International Telephone and Telegraph Consultative Committee (CCITT), now known as the International Telecommunications Union (ITU). For obvious reasons, ISDN uses a basic signal with a bandwidth of 64kbps. Though developed for speech transmission, the protocol also provides data channels of the same and smaller bandwidth. The Broadband version B-ISDN combines, by *multiplexing*, thirty 64kbps channels (and some signaling channels) to a bandwidth of 2Mbps. In the 1980s digital electrical networks provided higher and higher bandwidth, so that many signals could be multiplexed to a single signal. This technique is

called time division multiplexing (TDM) (or Time Division Multiple Access, TDMA). The principle is to divide time in frames, each of which is assigned to a lower bandwidth channel. In the *synchronous* version of TDM the channels are ordered and each channel gets fixed time slots. In figure 2 synchronous multiplexing is depicted: three OC1 are first multiplexed to one OC3 and then later demultiplexed to three OC1 again. The time frames are reduced with a factor 3 in this process, and empty frames are left empty.

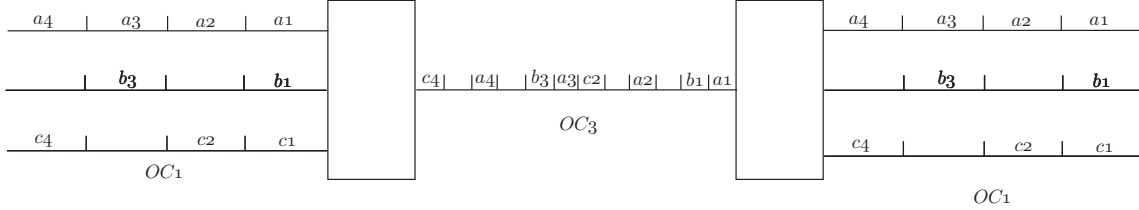


Figure 2: Synchronous Time Division Multiple Access

Digital networks opened the opportunity for mass transportation of data with high reliability. Data are generally not time critical, nor symmetric (data are sent from one source to a destination and generally not back). So, there is no need to claim a complete circuit. Packet switching (as opposed to circuit switching) was developed to make efficient use of this asymmetry. In packet switching data are partitioned into multiple packets each with a header containing information such as the size of the packet and the destination, where the packets are reassembled. In 1969 the development of packet switching technology started with the ARPANET (Advanced Research Projects Agency NETwork), a digital network connecting several American institutions. The protocol for ARPANET is the forerunner of TCP/IP the protocol for Internet. TCP/IP uses variable length packets (maximum size 64kbps, average size 1 kb) with destination address information. Each node in the network contains a router which decides on the basis of the receiver address and the shortest path to this address which link, incident to the router node, is selected. The shortest path is determined with routing protocols such as OSPF (Open Shortest Path First). In OSPF the network operator defines lengths on links to influence shortest paths between node pairs. This way, traffic can be regulated to avoid congested areas. The determination of optimal lengths is an interesting, but hard optimization problem. TCP/IP has no guaranteed quality of service. It assigns available bandwidth to a connection dynamically, based on packet-loss information. Thus, IP is unsuitable for time-critical transmission such as voice and video. This drawback has led to the development of ATM (Asynchronous Transmission Mode). ATM uses short fixed-length packets ( $48+5=53$  bytes) which are sent over dedicated virtual paths. A virtual path (VP) is a direct logical connection between a pair of vertices in the network of a certain capacity. The vertices need not be connected by a physical link in the network. VPs reduce the need for reading the packets in intermediate nodes (hubs). Nowadays virtual paths can be created dynamically based on demand, which ensures efficient, reliable usage of available bandwidth. ATM also allows *statistical multiplexing* by use of asynchronous TDM: packets from different data streams are combined over the same (virtual) connection. The price to pay is that the headers of the packets must be read at the endpoints of a link.

## Optical networks

IP and ATM are data transmission protocols with which a logical overlay network is created on top of a physical network. Physical networks have worked solely with electrical transmission over copper cables, until the end of the 1980s. Then transmission with light (through glass fiber) became available. With optical transmission higher speeds can be reached compared to electrical transmission. In the beginning 155Mbps (STM1) was the standard speed, currently 2.5Gbps (STM16) is implemented. Add-Drop Multiplexers (ADMs) combine and separate signals from STM1 to STM16 and back. These devices are responsible for the electrical-optical conversion. Speed increases are not the only means for capacity increase in optical systems. A fiber can also carry multiple wavelengths simultaneously (currently more than 100). Optical ADMs (OADM) combine and separate wavelengths. This is called (D)WDM, (Dense) Wavelength Division Multiplexing. This type of multiplexing adds a second dimension to multiplexing: Electromagnetic spectrum besides time. Another advantage of light over electricity is that it needs far fewer repeaters (amplifiers), and thus it is more reliable and cheaper to construct long-distance networks.

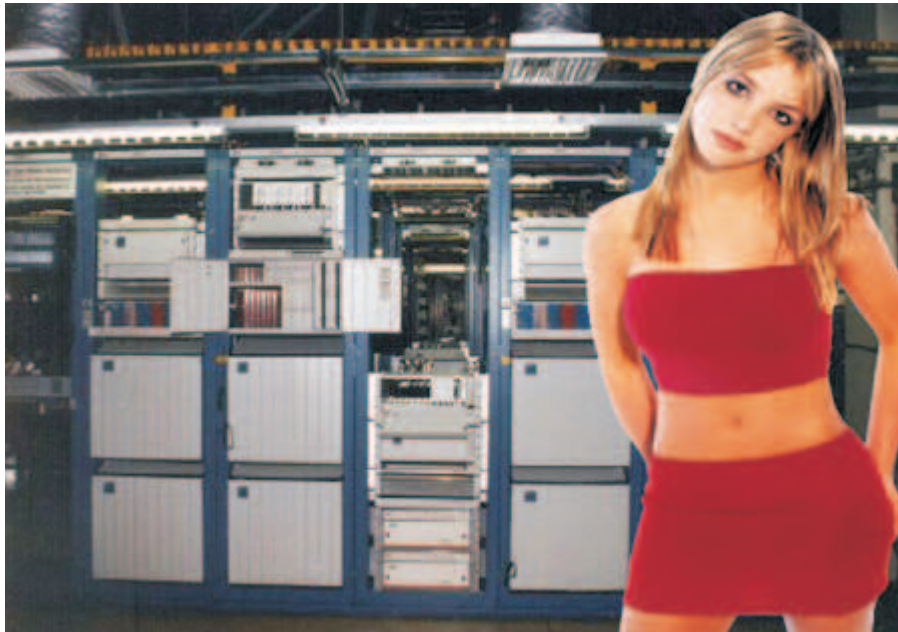


Figure 3: Optical switch equipment

A standard protocol for optical transmission is SDH/SONET (SDH = Synchronous Digital Hierarchy in Europe, SONET in the U.S.). With the SDH/SONET protocol the network developers started using rings as the basic network topology. These rings have a lot of advantages. They are simple in structure, and easy to set up with OADMs. Furthermore, protection and restoration of signals is easily implemented by using double fibers. There are two types of rings depending on the type of switching and the protection/restoration mechanism, both called self-healing because of the automatic and thus fast restoration capabilities: unidirectional path-switched rings (UPSR), and bidirectional line-switched rings (BLSR). In a UPSR, traffic is sent in one direction over one fiber, say clockwise, and a back-up signal is

sent on another fiber counter-clockwise. The back-up signal is automatically picked-up when the original signal is blocked. In a BLSR, a pair of fibers is used differently. Again traffic is sent in one direction over one fiber, say clockwise. However, the second fiber, the back-up fiber, is used as follows. If a link (or node) fails, all traffic that is blocked is redirected automatically over the back-up fiber. Currently, BLSRs are implemented with a double set of fiber pairs (four fibers in total), where one pair is used for clockwise routing and another pair for counter-clockwise routing. The four-fiber BLSR is more economical with capacity usage than the two-fiber BLSR or the UPSR, since (back-up) routes are on average shorter.

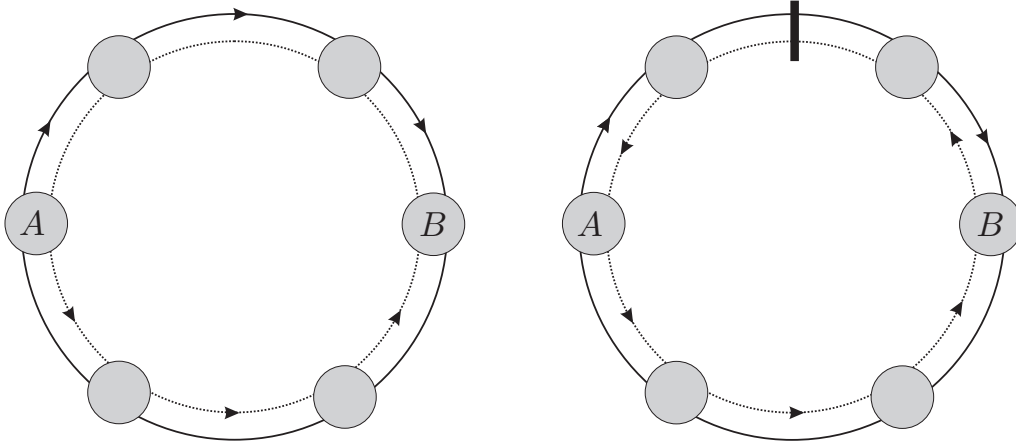


Figure 4: UPSR and BLSR

### 1.2.2 Wireless telecommunication

The idea of cell-based mobile radio systems as we know them by now, was developed at Bell Laboratories (in USA) in the early 1970s. An area, such as a city is divided in small hexagonal cells. In the center of each cell a base station is situated which communicates with mobile phones present in the cell, through a number of antennas. Each cell may contain multiple antennas for communication, depending on the demand. The base stations are connected through a fixed network. Cellular systems were first implemented in the USA with the release of AMPS (Advanced Mobile Phone Service) in 1983, an analog transmission system. The success of mobile telephony came in the early 1990s when digital communication was introduced. The US introduced the ADC system (American Digital Cellular), and Europe developed GSM (Global System for Mobile Communication). Though both systems are similar in structure and technology there are small differences. Both use compression techniques to reduce bandwidth for speech to about 10 kbps. GSM uses TDMA in combination FDMA. FDMA (open air) is similar to DWDM (in fibers), namely a division of the electromagnetic spectrum into small frequency bands. A consequence of open air communication is that the frequency bands should be selected carefully for antennas with a small geographical distance (in the same or neighboring cells), since interference of signals may cause unacceptable loss of quality. This problem makes frequency planning an important subject for GSM.

Wireless data transmission is a recent development. GPRS and its successor UMTS are

packet-switching protocols, which allow for high-bandwidth data transmission. UMTS uses channels with a fixed bandwidth of 5Mbps which is transmitted from a base station to all mobile stations. Each station transmits its information using Code Division Multiple Access (CDMA), i.e., a mobile phone selects only the data with the right code. Thus, the available bandwidth can be divided over the users according to their needs.

## 2 Classical network planning problems

Quality of Service (QoS) has always been the major issue for telecom providers: A customer is mainly interested in getting a good quality connection whenever he requests one. It is the telecom company's problem to have one ready. Clearly, it is too costly for the company to have a 100% guaranteed availability for all customers at any time. This is also not necessary since demand for calls or bandwidth capacity is varying over time: nobody always calls, nor do different calls always occur at the same time. So, the installed capacity is to have a guaranteed availability probability, the Quality of Service. Generally, the companies strive for a 99% connection guarantee at the busiest time of the day. In the early circuit-oriented telephone networks routes of calls were completely determined. The theory for determining the amount of capacity needed on a link in such networks, was developed in 1917 by Erlang (see [18]) and [19]). Erlang used queueing theory (Poisson processes) to determine the blocking probability of a link with a given capacity.

$$P(\rho, C) = \frac{\frac{\rho^C}{C!}}{\sum_{i=0}^C \frac{\rho^i}{i!}}$$

Here,  $\rho$  is the average number of simultaneous calls,  $C$  is the number of available lines, and  $P$  is the blocking probability: the proportion of callers whose calls are blocked. The availability requirements ( $P(\rho, C) < 0.01$ ) of the links of a network determine a minimum amount of capacity  $C$  on each of the links.

Nowadays, with digital transmission and packet-switching there are two problems: the routing of traffic between a given pair of nodes may vary, like in IP-networks, and data traffic is asymmetric and bursty. So the traffic demand on a link is much more unpredictable. Many studies have, however, revealed that the traffic has a stochastic behavior which is essentially independent of the volume of the traffic: *self-similarity* (see Crovella and Bestavros [15] and Leland et al. [54]).

The in-depth treatment of the stochastic processes modelling the determination of capacity demand is beyond the scope of this manuscript. We consider the operational processes involved in the efficient set-up and usage of a network. The three main components of these processes are:

1. Design: which links or edges to develop. Often to meet certain connectivity requirements for protection.
2. Synthesis: how much capacity to put on all links to serve all traffic demands.



3. Routing: which paths to use for the various traffic streams to meet demand without violating capacity restriction on links.

In this section we describe the standard models for design, synthesis and routing. We discuss their complexity and solution methods, and related fundamental mathematical results. The models are all strongly related to the multi-commodity flow problem. Routing is most closely related to the multi-commodity flow problem, and therefore treated first. Then Network Design and Synthesis are discussed.

## 2.1 Routing: multi-commodity flow models

Let  $G = (V, E)$  be an undirected network. Consider a set of commodities  $K$ , where each commodity  $k \in K$  is determined by a source-terminal pair  $(s_k, t_k)$  of vertices, and a demand  $d_k$ , which is to be routed over the network from  $s_k$  to  $t_k$ . Let  $u_{ij}$  represent an upper bound on the capacity of edge  $\{i, j\}$ . The problem is to decide how much demand is routed over each edge, without violating capacity constraints. The (linear) costs  $c_{ij}$  are determined by the actual capacity usage, and should be minimized. (For models allowing for nonlinear costs as well, we refer to Minoux [61]). The integer program of the problem uses the variables  $f_{ij}^k$ , indicating how much demand of commodity  $k \in K$  is routed from  $i$  to  $j$  over edge  $\{i, j\}$ . The complete formulation reads:

$$\min \quad \sum_{\{i,j\} \in E} \sum_{k \in K} c_{ij} (f_{ij}^k + f_{ji}^k) \quad (1)$$

$$\text{s.t.} \quad \sum_{j: \{i,j\} \in E} f_{ij}^k - \sum_{j: \{i,j\} \in E} f_{ji}^k = \begin{cases} d^k & \text{if } i = s_k \\ -d^k & \text{if } i = t_k \\ 0 & \text{otherwise} \end{cases} \quad \forall k \in K, \forall i \in V \quad (2)$$

$$\sum_{k \in K} (f_{ij}^k + f_{ji}^k) \leq u_{ij} \quad \forall \{i, j\} \in E \quad (3)$$

$$f_{ij}^k, f_{ji}^k \geq 0 \quad \forall k \in K, \forall \{i, j\} \in E \quad (4)$$

The constraints 2 are standard flow conservation constraints for each of the commodities. The capacity on an edge is *undirected* because installed capacity can be set-up for usage by traffic in both directions. Thus, the sum of forward and backward flow on an edge should not exceed its capacity. This is reflected by the constraints 3. The objective 1 may be used to set preference on shortest paths (in length or number of connections) although this is fairly artificial: feasibility is the important issue, and capacity availability rather than capacity usage determines the operational costs. Note that the multi-commodity flow problem is stated as a linear program here and, therefore, is polynomially solvable. In general the flows are restricted to be integral, as capacities come in standard units (such as STM16). The integer multi-commodity flow problem is NP-hard, already for instances with only two commodities (Even, Itai, Shamir [23] and Karp [44]). In telephone networks and ATM networks, the traffic

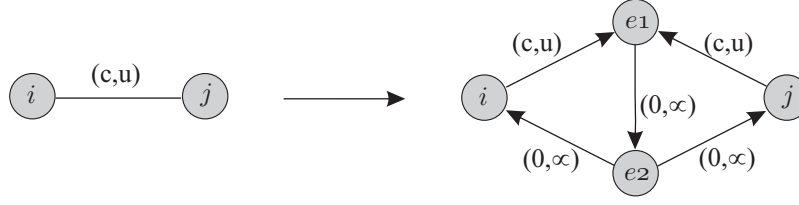


Figure 5: Transformation to directed multi-commodity flow

demand of a commodity can not be divided over multiple paths, but exactly one should be selected: non-bifurcated routing. The integer multi-commodity flow model can be adapted to this situation fairly easily: the variables  $f_{ij}^k$  are replaced by  $d_k y_{ij}^k$ , where the  $y_{ij}^k$  are binary variables.

A fundamental issue is the feasibility of the multi-commodity flow problem, both for the integer and linear version. An obviously necessary condition is the *cut condition*:

$$\sum_{\{i,j\} \in E: |\{i,j\} \cap S|=1} u_{ij} - \sum_{k \in K: |\{s_k, t_k\} \cap S|=1} d_k \geq 0 \quad \forall S \subset V \quad (5)$$

In the single commodity case, Ford and Fulkerson [26] showed that the cut condition is sufficient for feasibility: the MAXFLOW-MINCUT theorem. Though Ford en Fulkerson proved their result in directed networks it also applies to undirected networks, since these can be transformed to directed networks without additional commodities: Replace each edge  $e = \{i, j\}$  with the following construct. Two new nodes  $e1$  and  $e2$  are introduced. The arcs are  $(i, e1)$ ,  $(j, e2)$  with  $(\text{cost}, \text{capacity}) = (c_{ij}, u_{ij})$  and  $(e1, e2)$ ,  $(e2, i)$ ,  $(e2, j)$  with  $(\text{cost}, \text{capacity}) = (0, \infty)$ .

Hu [41] proves that the cut-condition is also sufficient for feasibility in case of two commodities, though the integral multi-commodity flow problem is NP-hard. Perhaps surprisingly, the cut condition is not sufficient for feasibility in case of more commodities, as the example in figure 6 shows.

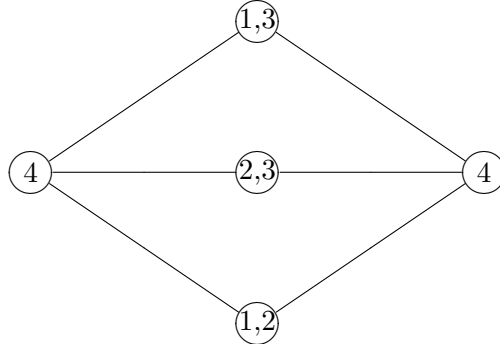


Figure 6: Counterexample for the sufficiency of the cut condition.

In this example there are four commodities, each with unit demand. The endvertices of the four commodities are given by the numbers in the nodes. The edges all have unit capacities. Each of the commodities 1, 2, 3 must use at least two units of capacity in the network, since the length of any path connecting the source and sink of a commodity is at least two. Since there are only six units available there is not enough for four commodities.

On the positive side the cut-condition is sufficient for specific networks and commodity structures. Okamura and Seymour [62] showed that, in planar networks, in case the endpoints of the commodities all lie on one face of the graph, the cut condition is indeed sufficient. For integral capacities and demands they showed that there is a half-integral flow under the given conditions. In fact, they prove a slightly stronger result: the integer version of the sufficient conditions requires that the left-hand side of 5 is even. Generalizations of the above conditions mainly relax the requirements for the commodities to lie on one face. See Frank [31] for an overview. Kramer and Van Leeuwen [50] showed that the integer multi-commodity flow problem remains NP-complete, even if the underlying graph is a grid, and if the capacities and demands are all unit. Thus, the position of the endvertices of the commodities is essential.

Though, the (fractional) multi-commodity flow problem is polynomially solvable with linear programming techniques, this is not a very practical method. A series of papers using combinatorial ideas has been published recently. These combinatorial ideas are applied on the dual of the LP in 1.4. They have led to fast and simple FPTAS (see Fleischer [25] and Leighton et al. [53]) and slightly more complicated exact methods such as described in Garg and Koenemann [34]. The solution methods for integer multi-commodity flow problems use standard techniques like Lagrangean relaxation (on the capacities) and Benders decomposition. An overview of these techniques is given in Assad [3].

## Path formulation

Instead of using flow variables on individual edges to model routing restrictions, one can also use flow variables associated with paths. Let  $z_p^k$  be the flow on a path  $p \in P_k$  (where  $P^k$  is the set of all simple paths connecting the endvertices of commodity  $k$ ). Let  $P_{ij}^k \subseteq P^k$  denote the set of paths for commodity  $k$  that contain the edge  $\{i, j\}$ . The path model reads

$$\min \quad \sum_{k \in K} \sum_{p \in P^k} c_p z_p^k \quad (6)$$

$$\text{s.t.} \quad \sum_{p \in P^k} z_p^k = d_k \quad \forall k \in K \quad (7)$$

$$\sum_{k \in K} \sum_{p \in P_{ij}^k} z_p^k \leq u_{ij} \quad \forall \{i, j\} \in E \quad (8)$$

$$z_p^k \geq 0 \quad \forall k \in K, \forall p \in P^k, \forall \{i, j\} \in E \quad (9)$$

The path model can be adapted to non-bifurcated routing in the same way as the flow model. In fact, this model is especially useful for non-bifurcated routing, in case restrictions specific for paths, such as a maximum length (hop restrictions), are to be incorporated. Such restrictions can not be handled in the flow model. The drawback of the path formulation

is the possibly huge amount of variables. Solving the integer multi-commodity flow problem with this formulation is therefore done using column generation techniques. See Ozdaglar and Bertsekas [63].

### Cut formulation

A truly undirected formulation can be obtained by converting the flow into a circulation. Here, we add for each commodity  $k$  the edge  $\{s_k, t_k\}$ , with capacity  $d_k$ . The flow  $f_{\{s_k, t_k\}}^k$  on  $\{s_k, t_k\}$  is set to the capacity  $d_k$ , so that no other commodities can make use of this edge. The flow variables  $f_{ij}$  now denote "undirected" flow on edge  $\{i, j\}$ .

$$\min \quad \sum_{\{i,j\} \in E} \sum_{k \in K} c_{ij} f_{ij}^k \quad (10)$$

$$\text{s.t.} \quad \sum_{e \in \delta(S) \setminus \{\bar{e}\}} f_e^k \geq f_{\bar{e}}^k \quad \forall k \in K, \forall S \subset V, \forall \bar{e} \in \delta(S) \quad (11)$$

$$\sum_{k \in K} f_{ij}^k \leq u_{ij} \quad \forall \{i, j\} \in E \quad (12)$$

$$f_{\{s_k, t_k\}}^k = d_k \quad \forall k \in K \quad (13)$$

$$f_{ij}^k \geq 0 \quad \forall k \in K, \forall \{i, j\} \in E \quad (14)$$

Here,  $\delta(S)$  is the set of edges with exactly one vertex in  $S$ . Constraints 11 express the condition that in any cut there is no edge containing more flow than all other cut-edges together. Seymour [65] has shown that the cut-constraints are necessary and sufficient to define a circulation. Bauer [8] extended this formulation to the integer case, where besides additional integrality constraints on the flow variables, an even-ness condition on the total flow on edges incident to any node  $v$  is added:

$$\sum_{e \in \delta(i)} f_e^k = 0 \pmod{2} \quad \forall i \in V, \forall k \in K \quad (15)$$

Note, that the constraints 15 are not linear. Recently, a cutting plane algorithm using the cut formulation has been described by Brunetta et al. [11]. They present valid inequalities based on the comb and clique-tree inequalities for the travelling salesman problem.

## 2.2 Network design: survivability with connectivity restrictions

The design of a network is steered by survivability requirements. Survivability is the network's ability to remain connected when network element failures occur. Generally, the survivability is measured in connectivity restrictions of pairs of vertices. That is, each pair of vertices  $s_k$  and  $t_k$  ( $k \in K$ ) requires a minimum number  $d_k$  of edge (or node) disjoint paths in the

network. Often, being able to recover from one failure suffices, since repair times are generally so small that the probability for a second failure to happen simultaneously, is negligible. Thus,  $d_k \in \{0, 1, 2\}$ . Often, connectivity requirements are expressed in node requirements  $r_i$  ( $i \in V$ ), and the connectivity requirement of a pair  $(s_k, t_k)$  is then the minimum of their individual requirements:  $d_k = \min\{r_{s_k}, r_{t_k}\}$ . Special cases are the minimum spanning tree problem: 1-connectivity for all nodes (pairs), and the Steiner tree problem: 1-connectivity for a subset of the nodes. Note that the problem is NP-hard, since the Steiner tree problem is a special case.

The flow formulation of the network design problem is closely related to the formulation of the multi-commodity flow problem. The major difference is the replacement of the capacity restrictions by the restrictions 18 that enforce an edge to be installed once it contains flow. Note that the objective 16 only contains the installation costs  $b_{ij}$  for the edges, and no costs for the flow.

$$\min \quad \sum_{\{i,j\} \in E} b_{ij} x_{ij} \quad (16)$$

$$\text{s.t.} \quad \sum_j f_{ij}^k - \sum_j f_{ji}^k = \begin{cases} d_k & \text{if } i = s_k \\ -d_k & \text{if } i = t_k \\ 0 & \text{otherwise} \end{cases} \quad \forall k \in K, \forall i \in V \quad (17)$$

$$\max\{f_{ij}^k, f_{ji}^k\} \leq x_{ij} \quad \forall k \in K, \forall \{i, j\} \in E \quad (18)$$

$$f_{ij}^k, f_{ji}^k \in \{0, 1\}, x_{ij} \in \{0, 1\} \quad \forall k \in K, \forall \{i, j\} \in E \quad (19)$$

The following model is obtained by projection of the flow formulation 16-19 onto the space of the design variables  $(x_{ij})$ . The correctness of this formulation follows from Menger's theorem [60]: The number of edge-disjoint paths between two vertices  $s$  and  $t$  is the minimum number of edges in a cut separating  $s$  and  $t$ .

$$\min \quad \sum_{\{i,j\} \in E} b_{ij} x_{ij} \quad (20)$$

$$\text{s.t.} \quad \sum_{\{i,j\} \in \delta(S)} x_{ij} \geq \text{con}(S) \quad \forall S \subset V, S \neq \emptyset, V \quad (21)$$

$$x_{ij} \in Z_0^+ \quad \forall k \in K, \forall \{i, j\} \in E \quad (22)$$

Here  $\text{con}(S)$  is the maximum connectivity requirement over the pairs of vertices  $(s_k, t_k)$  separated by  $S$ , i.e., having exactly one of the two vertices in  $S$ .

The LP-relaxation of both formulations is quite weak. Various types of partition inequalities have been developed to strengthen the LP-relaxation. See Grötschel et al. [38] and Mahjoub [59]. However, the use of them in Branch-and-Cut algorithms has had limited success. Magnanti and Raghavan [58] introduce a stronger directed network version based on a result of Nash-Williams: The edges of a graph can be directed such that for each pair of

vertices  $s$  and  $t$  the maximum number of edge-disjoint paths can be split into two halves: one half of the paths (rounded down) moves from  $s$  to  $t$  and the other half from  $t$  to  $s$ . This result admits a strengthened version of the cut inequalities. In Gabow et al. [33] a  $2k - 1$  approximation algorithm is given for the problem to design a minimum cost  $k$ -connected network.

### 2.3 Simultaneous network synthesis: loading capacity

The Network Synthesis problem, differs in only one way from the multi-commodity flow problem. Capacity on each of the edges is not available, but should be installed at a certain (unit) cost. Thus, like in the design problem, we have besides the flow variables  $f_{ij}^k$  a set of extra variables, the design variables  $x_{ij} \in Z_0^+$  denoting the capacity installed on edge  $\{i, j\}$ . Let  $b_{ij}$  ( $\{i, j\} \in E$ ) represent the unit installation costs of capacity. The objective is to minimize the total installation costs. Flow costs are generally not taken into account. The capacity constraints in the multi-commodity flow model are replaced by the following constraints.

$$\sum_{k \in K} (f_{ij}^k + f_{ji}^k) \leq x_{ij} \quad \forall k \in K, \forall \{i, j\} \in E \quad (23)$$

The capacities generally come in units, so the design variables are forced to be integral. If the capacities were not restricted to be integral the problem would be a linear program. Johnson, Lenstra, and Rinnooy Kan [43] formulate the problem with a budget on capacity installation costs, and show that this problem is NP-hard. In fact, it could be separated into  $|K|$  shortest path problems, one for each commodity. Solution techniques for the problem have mainly used the polyhedral approach. In Magnanti et al. [56] and [57] several classes of valid inequalities are used in a Branch-and-Cut approach. A large class of inequalities are the cut-inequalities, which are derived from the cut-condition 5 by replacing the capacity parameters  $u_{ij}$  by the capacity variables  $x_{ij}$ . Many other inequalities are based on the substructure of a single edge. This particular subproblem has been studied by Atamtürk [5] and van Hoesel et al. [40]. A more general version of the capacity constraints occurs when the capacity is available in two (or more) sizes, say 1 and  $C$  (Currently, STM1 and STM16 are popular sizes). Then the RHS of 23 is replaced by  $x_{ij} + Cy_{ij}$ , where  $x_{ij}$  measures the number of units of size 1, and  $y_{ij}$  measures the number of units of size  $C$ . Note that the cost of installing a unit of capacity  $C$  should be less than  $Cb_{ij}$ , for otherwise it will be cheaper to install  $C$  separate units. Magnanti et al. [57] derive cut-set inequalities for this problem, which have been the subject of many later studies such as Bienstock and Günlük [9] and Atamtürk [4].

An old special case occurs in complete graphs, if the flow costs are zero, and capacity installation costs are equal on all links. Gomory and Hu [37] developed a polynomial-time two-phase procedure that determines the minimum total load: First, a maximum capacity tree is determined using the demands as edge-weights; second, the tree is divided in a set of maximal subtrees with uniformly distributed capacities and each of these subtrees is replaced

by a cycle through all nodes of the tree, with cycle-capacity being half of the subtree capacity. See Minoux [61] for a detailed description.

## 2.4 Non-simultaneous network synthesis: Back-up capacity

Back-up networks have spare capacity that can be used if demand exceeds available capacity in the main network. Generally, demand excess will occur for few commodities simultaneously at any point in time. In case we wish to protect a network against demand excess for single commodities we get non-simultaneous network synthesis problem. Here, for a given set of excess demands capacity is to be placed such that each of the demands can be routed individually, not simultaneously. A formulation of this problem is obtained from the network design problem by replacing the binary restrictions on the design variable by integrality restrictions

A notable special case occurs in complete graphs, if capacity installation costs are equal for all links, and where the edges to be selected are restricted to form a tree. The Gomory-Hu tree [36], determines the optimal loads as follows. The tree is computed by taking sequentially a series of minimum cuts between pairs of vertices. These pairs are selected such that they are not separated by previously found cuts. Gomory and Hu show in [36] that these cuts are non-crossing, i.e., for two pairs of cuts defined by the subsets  $S \subset V$  and  $T \subset V$  either  $S$  or its complement  $\bar{S}$  is contained in  $T$  or  $\bar{T}$ . This property ensures that after  $|V| - 1$  iterations each pair of vertices is separated by at least one cut. The next step is to connect the vertices that are separated by exactly one cut. The connecting edge gets the weight of this cut (the minimum cut). The thus formed graph is a tree with the following property: the minimum capacity cut between any pair of vertices  $v$  and  $w$  has a value equal to the minimum weight over the edges on the path connecting  $v$  and  $w$ . For each edge  $\{v, w\}$  we denote by its capacity the capacity demand of the commodity with endvertices  $v$  and  $w$ . Now, installing capacities on the edges of the Gomory-Hu tree equal to the corresponding cut capacities is optimal in the sense that we use a minimum amount of capacity to satisfy the non-simultaneous demand.

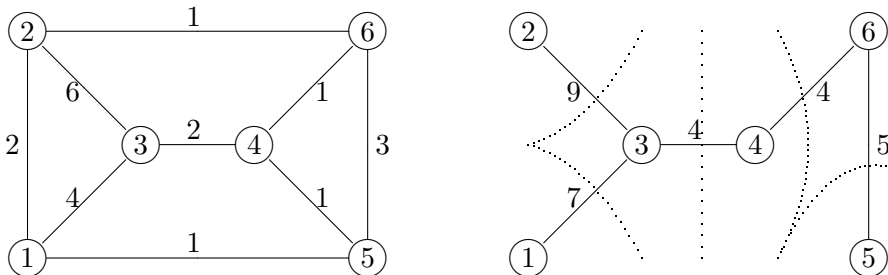


Figure 7: Example Gomory Hu tree.

The Gomory-Hu tree can also be used to determine the minimum cut between any pair of vertices: Take the path between the vertices in the tree and select the edge with the smallest weight.

Most studies restrict the set of edges on which capacity can be installed to form a tree. Then the problem is the so-called communication spanning tree problem. Hu [42] gave a weak

condition under which the optimal tree is a star. Wu et al. [72] use the star as the basic ingredient to develop an  $O(\log n \log \log n)$  approximation algorithm. Fischetti et al. [24] develop a Branch-and-Bound algorithm. In case all demands are unit, the problem is known as the minimum routing cost tree problem. Wu et al. [72] developed a FPTAS by using the condition of Hu [42]. Wu et al. create the tree with stars to limit the depth of the tree. Earlier, with other techniques, Wong [71] derived a 2-approximation. This special case is already NP-hard, see Johnson et al. [43].

## 2.5 Protection and restoration

Connectivity of a network is generally not enough to protect against failure of network elements. If a defective edge is detected, then the paths that use this edge should be restored by rerouting as quickly as possible. Various schemes are deployed to realize this.

The 1+1 protection mechanism dedicates two paths for each commodity: The working path that operates under normal conditions, and a reserve path with the same capacity that is edge disjoint with the working path. Clearly, 1+1 protection is capacity intensive. To minimize bandwidth installation costs, the problem is to find a cycle through source and sink of the commodity with minimum length. The high bandwidth demand of 1+1 protection has made network planners look for other protection mechanisms in which protection paths share capacity, the 1:N protection mechanism, where reserve capacity on each edge can be used by multiple protection paths. Optimization models to minimize spare or both working and spare capacity are quite complicated. It involves, for each commodity  $k$ , the determination of a working path  $P_k$  and a protection path  $Q_k$  which are edge disjoint. The determination of the necessary spare capacity on each edge  $e$  is the complicating factor:

$$\max_{f \in E-e} \sum_{k: e \in Q_k, f \in P_k} d_k$$

ILP models for the minimization of shared capacity can be found in Van de Leensel [52].

A way of shared path protection is link protection, where a dedicated path connecting the vertices of an edge is determined for rerouting all paths using the edge in case of failure. The advantage of this method is that it is fairly easy to implement in the network, though capacity consumption is fairly high. Link protection can be modelled as a non-simultaneous demand problem. Diversification can be used in case of bifurcated routing. Here, multiple edge disjoint paths are used in the original network each with capacity installed which is a fraction of the actual demand. If an edge fails at most one of the paths becomes unusable. So, installing a little extra capacity on all paths ensures the communication needs of the commodities. Besides limited extra capacity, diversification copes with the problem of signalling in a trivial way: There is no need for provisioning signalling, since the capacity is available on paths that are already operating. This idea is used in Koster et al. [49] to determine optimal capacity usage for diverse routing.



### 3 New network planning problems: virtual topologies

Virtual topologies and Wavelength Division Multiplexing (WDM) have introduced new problems in all three traditional areas of telecom network planning: routing, design and synthesis. These problems are the topic of this section. Virtual paths have been introduced in ATM networks for reliability reasons. The development of Add-Drop Multiplexers (ADMs) has enabled the construction of virtual paths in optical networks as well. An ADM is placed in a node of the network where it connects two fibers: One inbound and one outbound. Let us consider an STM16 signal over the fiber. The ADM can select any subset of the 16 incoming STM1 channels and convert them into electrical signals. Moreover, the freed channels can be replaced by other channels. In WDM the first step is to separate and/or combine wavelengths with Optical ADMs which operate in a manner similar to standard ADMs.

In the sequel we concentrate on virtual paths in WDM technology: *lightpaths*. Consider a physical network  $G = (V, E)$  and a set of point-to-point demands. The construction of virtual paths in the network follows three stages:

1. **Virtual Topology Selection.** Determine the edges of the virtual network, i.e., the lightpaths in terms of their source and destination, not the physical paths connecting the source and destination.
2. **Routing and Wavelength Assignment (RWA).** Determine the physical paths in the network of the lightpaths, and assign a wavelength to them. Paths that use the same fiber (on the same link) in the physical network must be assigned different wavelengths. Moreover, there is a limit on the number of wavelengths that can be used on fibers. The current maximum is about 100, but the actual number depends on the equipment.
3. **Traffic Routing or Grooming.** The actual traffic demands must be assigned a path in the virtual network. Grooming is the process of merging low bandwidth traffic streams over lightpaths, using (Optical) ADMs.

We discuss these stages on different physical/logical topologies, such as trees, rings, and general networks. The (self-healing) ring is a very popular structure in optical networks. Therefore we devote two special sections to it, one for the ring loading problem and one for the ring design problem.

#### 3.1 Routing and wavelength assignment

Consider a network  $G = (V, E)$  in which a set of vertex-pairs is given. Each pair is to be assigned a (light)path. Different lightpaths using the same fiber, must be assigned different wavelengths (colors). Thus, there are two things to be determined: the path of a node-pair and its color. The first version of the problem we consider, is where the paths are predetermined, i.e., they are part of the input. Thus, in  $G$  a set of (di)paths  $\mathcal{P} = P_1, \dots, P_K$  is given, for instance the shortest paths between the vertices of each pair. The problem is to assign colors (wavelengths) to these paths such that paths sharing an edge (in the same direction) use

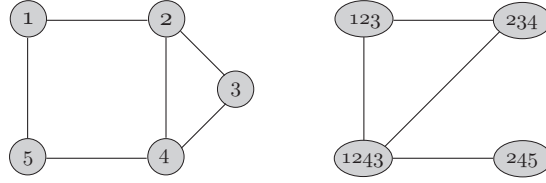


Figure 8: Construction of the conflict graph.

different colors. The objective is to minimize the number of colors used. This is the path-coloring problem. The problem is *symmetric* if the paths are undirected, and *asymmetric* if the paths are directed. The combined routing and coloring problem, is discussed after the coloring problem.

The path-coloring problem has a very natural transformation to vertex coloring: create a graph  $H = (W, F)$ , where the node set  $W$  consists of the (di)paths, and the edges are created by comparing paths: if two paths have an edge in common in  $G$ , then their corresponding nodes in  $H$  are connected.  $H$  is called the conflict graph. Note that for directed paths the direction on the edge should also be the same.

Consider the graph  $G$  on the left in figure 8, and the four paths  $P_1 = \{1, 2, 3\}$ ,  $P_2 = \{1, 2, 4, 3\}$ ,  $P_3 = \{2, 3, 4\}$ ,  $P_4 = \{2, 4, 5\}$  in  $G$ . The conflict graph of these paths is the graph on the right.

In Erlebach and Jansen [20] the complexity question of path-coloring is discussed. Both the symmetric and asymmetric problems are easy on chain graphs (graphs consisting of a single path). For the symmetric problem the conflict graph is an interval graph. The asymmetric problem is twice a symmetric problem, one for each direction on the chain. Erlebach and Jansen [20] also show that the asymmetric coloring problem on binary trees is NP-hard. The symmetric problem is equivalent to the arc-coloring problem, and therefore NP-hard on trees, but polynomially solvable on trees with bounded degrees of the vertices, which includes the binary tree case. Finally, both problems are NP-hard on rings.

On trees a  $\frac{5}{3}$ -approximation algorithm is known, and there is no approximation algorithm with ratio better than  $\frac{4}{3}$ , for both the symmetric and asymmetric problem. On rings there is a trivial 2-approximation: split the ring at an arbitrary vertex  $v$  and a chain coloring problem remains. The conflicting colorings of the paths passing  $v$  can be resolved by using at most the same amount of colors. In general graphs, no approximation algorithm with a ratio better than 2 exists. See Erlebach and Jansen [20].

The combined routing-and-coloring problem on trees reduces to the coloring problem on trees. On general graphs an ILP formulation based on the multi-commodity flow problem is known: we introduce additional parameters for the wavelengths available. In fact, we make a copy of the network for each wavelength. The following formulation, from Zang et al. [73] minimizes the maximum load in the network.

$$\min \quad f_{max} \tag{24}$$

$$\text{s.t.} \quad \sum_j f_{ij}^{kw} - \sum_j f_{ji}^{kw} = \begin{cases} d^{kw} & \text{if } i = s^k \\ -d^{kw} & \text{if } i = t^k \\ 0 & \text{otherwise} \end{cases} \quad \forall k \in K, \forall i \in V \quad (25)$$

$$\sum_{k \in K} \sum_{w \in W} f_{ij}^{kw} \leq f_{max} \quad \forall \{i, j\} \in E \quad (26)$$

$$\sum_{w \in W} d^{kw} = d^k \quad \forall k \in K \quad (27)$$

$$\sum_{k \in K} f_{ij}^{kw} \leq 1 \quad \forall \{i, j\} \in E \quad (28)$$

$$f_{ij}^{kw} \in \{0, 1\} \quad \forall k \in K, w \in W \quad (29)$$

Heuristics have been described by Zang et al. [73] and Kennington et al. [45] among others. An overview and some more results can be found in Auletta et al. [6] and in Caragiannis et al. [12]. Hassin and Levin [39] treat the version where the number of colors available on an edge is fixed to  $C$ , minimizing the number of edges: the synthesis version of the problem. They prove that this problem is NP-hard (for  $C \geq 2$ ), and they give a  $2\frac{1}{2}$  approximation algorithm.

Wavelength Converters (WCs) have the capability to change the wavelength of an incoming lightpath at a node to a different wavelength of the leaving path. Thus, a path is split into two separately colorable paths. Thus, converters can be used to decrease the number of necessary wavelengths on links, by placing them on suitable nodes in a network. Wavelength converters are a dominant cost factor. So, the interesting optimization question is how to position them so as to minimize their use, and simultaneously reduce the need for different wavelengths. We first consider converters that can change many wavelengths simultaneously in a network node.

Consider a graph  $G = (V, E)$  and a set of paths  $\mathcal{P} = P_1, \dots, P_k$ . The load of an edge is defined as the number of paths that use the edge. The maximum load over all edges is a lower bound on the necessary number of wavelengths, and therefore an important measure for wavelength usage. Since paths need not be fixed forever in a network and demand can vary as well over time, Wilfong and Winkler [70] consider all possible sets of paths in  $G$  and ask themselves the question how to position converters on a subset  $S$  of the nodes, such that for any set of paths  $\mathcal{P}$  the number of necessary wavelengths is equal to the maximum load defined by  $\mathcal{P}$ . Such a set is called a *sufficient set*. The Minimum Sufficient Set problem (MSS) is to find the smallest sufficient set  $S$ . Again we have the symmetric and the asymmetric version of the problem. The asymmetric version has been shown to be NP-hard already for planar networks by Wilfong and Winkler [70]. Kleinberg and Kumar [47] give a 2-approximation algorithm on general graphs. They also provide a polynomial approximation scheme for planar graphs. This result has been extended to general graphs by Erlebach and Stefanakos [22] who also give a linear algorithm if the underlying network has bounded tree-width. A *spider* is a tree that consists of a central node and paths emerging from it. Wilfong and Winkler [70] have identified spiders as the only graph class that has an MSS of size 0, i.e., a spider does not need any wavelength converters. (Rings have an MSS of 1.) The placement of a WC has the effect of separating all edges incident to a node  $v$  by replacing  $v$  with a new node for each edge. Now, a sufficient set has the property that this copying operation converts the graph

into a set of spiders.

The symmetric problem is easy. This follows directly from the characterization of graphs with MSS of size 0: paths. Thus, WCs must be placed on all nodes with degree  $\geq 3$ . The only exception are rings, where an arbitrary node of degree 2 must be selected. Erlebach and Stefanakos [21] restrict themselves to sets of paths  $\mathcal{P}$  which contain shortest paths only, bearing in mind that these paths are the ones generally used in network routings. They characterize graphs with MSS of size 0. A somewhat surprising result can be found in Koster and Zymolka [48]. They consider WCs that can handle one wavelength change only. For a given set of paths (part of the input), they show that the placement of WCs is NP-hard already in star networks. They also provide heuristics for their version of the problem.

### 3.2 Grooming

The grooming problem is a multi-commodity flow problem once the virtual topology is determined. Therefore, the grooming problem is generally combined with the Virtual Topology Selection problem, such that the wavelength assigned is ignored. The ratio behind this is that wavelengths are sufficiently available (or will be in the future), and the costs of building a virtual network are determined by the ADMs. As each lightpath needs an ADM at its endvertices the minimization of lightpaths is closely related to the minimization of ADMs.

The problem of minimizing the number of ADMs is NP-hard, even for networks restricted to paths and stars. See Dutta et al. [17]. Some approximation results have been achieved by Calinescu and Wan [32]. They show that the minimum number of ADMs on a ring can be approximated within a ratio of  $\frac{3}{2}$ .

Good grooming solutions may not only minimize the number of ADMs used, but may simultaneously use a small number of wavelengths. The following example from Chow and Lin [13] shows, however, that both objectives can not always be optimized simultaneously. Consider a bidirectional ring of nine vertices (numbered 1..9) with unit symmetric demand between the vertices 1,2,3 and 4,5,6 and 7,8,9. Each wavelength has unit capacity in both directions. Then a wavelength minimizing solution uses two wavelengths, one with demands 12, 23, 45, 56, 78, 89 and one with demands 13, 46, and 79 for a total of 15 (9+6) ADMs. Another solution uses three wavelengths, one with demands 12, 23, 13, one with demands 45, 56, 46, and one with demands 78, 89, 79, using a total of 9 ADMs.

The following formulation solves the combined topology and grooming problem on UPSRs. In Sutter et al. [67] this formulation is introduced, though SDH/SONET rings were used instead of wavelengths. Note that from a mathematical point of view they are equal. The objective is determined by the number of (O)ADMs, which is equivalent to the number of nodes on the rings. The following model optimizes the number of used (O)ADMs. Flow variables  $f_{s_k, t_k}^r$  denote the traffic of commodity  $k$  assigned to ring  $r$ . The variables  $x_i^r$  are set to one, if ring  $r$  contains an ADM at node  $i$ .

$$\min \quad \sum_{r \in R} \sum_{i \in V} x_i^r \quad (30)$$

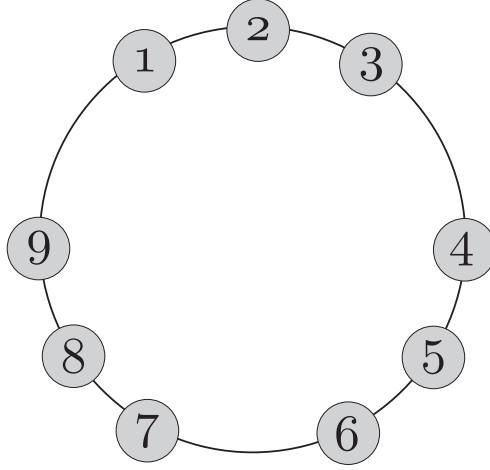


Figure 9: Example grooming.

$$\text{s.t.} \quad \sum_{r \in R} f_{s_k, t_k}^r = d^k \quad \forall k \in K \quad (31)$$

$$f_{s_k, t_k}^r \leq d_k \min(x_{s_k}^r, x_{t_k}^r) \quad \forall r \in R, \forall k \in K \quad (32)$$

$$\sum_{k \in K} f_{s_k, t_k}^r \leq C \quad \forall r \in R \quad (33)$$

$$f_{s_k, t_k}^r \geq 0 \quad \forall k \in K, \forall r \in R \quad (34)$$

$$x_i^r \in \{0, 1\} \quad \forall r \in R, \forall \{i, j\} \in E \quad (35)$$

If the traffic is not allowed to be split (non-bifurcated) we can replace each flow variable with the product of demand and a binary variable, see Sutter et al. [67]. In [67] the problem is then reformulated as a column generation problem and solved. In Sherali et al. [66] the problem allowing bifurcation is solved with a Branch-and-Cut procedure. Finally, the general grooming problem, on a mesh (complete graph), is solved heuristically by Battiti and Brunato in [7] and [10].

### 3.3 Ring loading: routing and grooming on 4 fiber BDSRs

In bidirectional self-healing rings the routing of the traffic demands can be done in two ways: clockwise and counter-clockwise. The optimization problem is to route the traffic such that the maximum capacity usage over all edges is as small as possible. This is reflected in the following formulation.

$$\min \quad \max_{k \in \{0, \dots, n-1\}} L_k \quad (36)$$

$$\text{s.t.} \quad L_k = \sum_{[i,j]: k \in [i,j]} \alpha_{i,j} d_{i,j} + \sum_{[i,j]: k \notin [i,j]} (1 - \alpha_{i,j}) d_{i,j} \quad \forall k \in \{0, \dots, n-1\} \quad (37)$$

$$\alpha_{i,j} \in \{0, 1\} \quad \forall \{i, j\} \in E \quad (38)$$

Here, the vertices and edges of the ring are numbered from 0 to  $n-1$ . Edge  $k$  is given by the vertices  $k$  and  $k+1$ , all modulo  $n$ . The interval  $[i, j)$  refers to the clockwise path from vertex  $i$  to vertex  $j$ , i.e., the edges  $i, \dots, j-1$ . The variables  $\alpha_{i,j}$  determine whether the clockwise path  $\alpha_{i,j} = 1$  or the counter-clockwise path  $\alpha_{i,j} = 0$  is used. The constraints 37 determine the load  $L_k$  on link  $k$ .

The problem originates from Cosares and Saniee [14], who proved that it is NP-complete, and who gave a 2-approximation algorithm. To show NP-completeness, a straightforward reduction from 2-PARTITION is given in Schrijver et al. [64]. If traffic can be bifurcated, i.e. the variables are relaxed to the interval  $[0, 1]$ , then the problem turns into a polynomial-size linear program. Note that each commodity has only two paths available.

The cut-condition 5 of Okamura and Seymour [62] is applied in Schrijver et al. [64] as well as in Dell'Amico et al. [16] for solving the LP-relaxation with a "greedy" algorithm. Rounding techniques led to an absolute error of  $\frac{3}{2}d_{max}$  guarantee in Schrijver et al. [64], where  $d_{max}$  is the largest demand. The latter then prove with their algorithm that integer bifurcated routing can be done in polynomial time. Thus, the problem with unit demands can be solved polynomially, as was proved first by Frank [30]. Dell'Amico et al. [16] use their algorithm in a Branch-and-Bound scheme to solve the problem to optimality. Finally, Khanna [46] develops a polynomial time approximation scheme using the greedy algorithm in Schrijver et al. [64].

### 3.4 Ring design

Many providers of optical networks use the ring structure as the basic element to create networks. In practice there are two ways of designing these rings: dense and sparse. In dense ring topologies any pair of nodes in a network is connected by at least one ring. The Dutch network of KPN provides an example. In sparse networks rings are connected to one another with an optical cross-connect (OXC). The OXCs are very expensive and therefore generally only one or two nodes are shared by a pair of rings. An example of a sparse network is the European backbone of many international operators.

For the ring design problem, we consider unidirectional rings. The demand routed over a unidirectional ring is equal to the capacity needed. The objective is to select a set of rings to satisfy demand, such that the number of (O)ADMs is minimized. Since each vertex of a unidirectional ring has an (O)ADMs, this is equivalent to the minimizing the number of nodes over all rings. The following model uses flow variables  $f_{s_k, t_k}^r$  which denote the traffic of commodity  $k$  assigned to ring  $r$ . The variables  $x_i^r$  are set to one, if ring  $r$  contains node  $i$ .

$$\min \quad \sum_{r \in R} \sum_{i \in V} x_i^r \quad (39)$$

$$\text{s.t.} \quad \sum_{r \in R} f_{s_k, t_k}^r = d^k \quad \forall k \in K \quad (40)$$

$$f_{s_k, t_k}^r \leq d_k \min(x_{s_k}^r, x_{t_k}^r) \quad \forall k \in K, \forall r \in R \quad (41)$$

$$\sum_{k \in K} f_{s_k, t_k}^r \leq C \quad \forall r \in R \quad (42)$$

$$f_{s_k, t_k}^r \geq 0 \quad \forall k \in K, \forall \{i, j\} \in E \quad (43)$$

$$x_i^r \in \{0, 1\} \quad \forall k \in K, \forall \{i, j\} \in E \quad (44)$$

The related model for non-bifurcated flow is found in Sutter et al. [67]. In [67] the problem is reformulated as a column generation problem and solved this way. In Sherali et al. [66] the (scaled) problem allowing bifurcation is solved with a Branch-and-Cut procedure.

Laguna [51] treats the sparse case. Laguna models the problem, and solves it heuristically with tabu search. The paper of Goldschmidt et al. [35] treats a specific structure: a central ring with other rings attached to it. It also describes an integer program and introduces heuristics to solve the problem. Finally, the Fortz et al. [28] describe the ring design problem with specific side-constraints such as a hop limit, and connectivity requirements of the rings: such as dual homing.

## 4 Concluding remarks

The previous sections contain a historical overview and quite recent results in network planning problems for telecommunication. By no means this overview is complete. For instance hop-constraints have not been treated. Moreover, broad classes of problems have not found a place here. This involves routing in IP networks as treated in Fortz et al. [29] and [27]. Routing in pricing problems is treated in, among others, Bouhtou et al. [55] and [69], and an interesting variant of the network design problem: the winner determination problem, where many customers compete for capacity in a network. This subject is found in Abrache et al. [2] and van Hoesel and Müller [68]. Finally, the extensive literature on frequency assignment in wireless networks has not been treated. A recent and quite complete overview can be found in Aardal et al. [1].

## References

- [1] K.I. Aardal, C.P.M. van Hoesel, A.M.C.A. Koster, C. Mannino, and A. Sassano. Models and solution techniques for the frequency assignment problem. *4OR*, 1, 2003.
- [2] J. Abrache, T.G. Crainic, and M. Gendreau. Design issues for combinatorial auctions. *4OR*, 2(1):1–33, 2004.
- [3] A. Assad. Multicommodity network flows-a survey. *Networks*, 8:37–91, 1978.
- [4] A. Atamturk and D. Rajan. On capacitated network design cut-set polyhedra. *"Mathematical Programming"*, 92(3):425–437, 2002.
- [5] A. Atamturk and D. Rajan. "on splittable and unsplittable capacitated network design arc-set polyhedra. *"Mathematical Programming"*, 92(2):315–333, 2002.

- [6] V. Auletta, I. Caragiannis, L. Gargano, C. Kaklamanis, and P. Persiano. Sparse and limited wavelength conversion in all-optical tree networks. *"Theoretical Computer Science"*, 266:887–934, 2001.
- [7] R. Battiti and M. Brunato. Reactive search for traffic grooming in WDM networks. *Lecture Notes in Computer Science*, 2170:56–64, 2001.
- [8] P. Bauer. The circuit polytope: facets. *Mathematics of Operations Research*, 22:140–145, 1997.
- [9] D. Bienstock and O. Günlük. "Capacitated network design — polyhedral structure and computation". Working paper, June 1995.
- [10] M. Brunato and R. Battiti. A multistart randomized greedy algorithm for traffic grooming on mesh logical topologies, 2002.
- [11] L. Brunetta, M. Conforti, and M. Fischetti. A polyhedral approach to an integer multi-commodity flow problem. *Discrete Applied Mathematics*, 101:13–36, 2000.
- [12] I. Caragiannis, C. Kaklamanis, and P. Persiano. Wavelength routing in all-optical tree networks: A survey. *"Computers and Informatics"*, 20:95–120, 2001.
- [13] T. Chow and P. Lin. "The ring grooming problem". *Networks*, 44:194–202, 2004.
- [14] S. Cosares and I. Saniee. An optimization problem related to balancing loads on sonet rings. *Telecommunications Systems*, 3:165–181, 1994.
- [15] M.E. Crovella and A. Bestavros. Self-similarity in world wide web traffic: Evidence and possible cause. *IEEE/ACM Transactions on Networking*, 6:835–846, 1977.
- [16] M. Dell'Amico, M. Labbé, and F. Maffioli. Exact solution of the sonet ring loading problem. *Operations Research Letters*, 25:119–129, 1999.
- [17] R. Dutta, S. Huang, and G.N. Rouskas. "Traffic grooming in path, star, and tree networks: Complexity, bounds, and algorithms". Working paper, North-Carolina State University, 2003.
- [18] A.K. Erlang. The theory of probabilities and telephone conversations. *Nyt Tidsskrift for Matematik B*, 20, 1909.
- [19] A.K. Erlang. Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Elektroteknikerne*, 13, 1917.
- [20] T. Erlebach and K. Jansen. The complexity of path coloring and call scheduling. *Theoretical Computer Science*, 255(1-2):119–137, 2001.
- [21] T. Erlebach and S. Stefanakos. On shortest path all-optical networks without wavelength conversion requirements. Tik report 153, Institut für Technische Informatik und Kommunikationsnetze, ETH Zurich, 2002.
- [22] T. Erlebach and S. Stefanakos. Wavelength conversion in networks with bounded treewidth. Tik report 132, Institut für Technische Informatik und Kommunikationsnetze, ETH Zurich, 2002.



- [23] S. Even, A. Itai, and A. Shamir. On the complexity of timetable and multicommodity flow problems. *SIAM Journal of Computing*, 5:691–703, 1976.
- [24] M. Fischetti, G. Lancia, and P. Serafini. Exact algorithms for minimum routing cost trees. *Networks*, 39(3):161–173, 2002.
- [25] L. Fleischer. Approximating fractional multicommodity flows independent of the number of commodities. *SIAM Journal of Discrete Mathematics*, 13(4):505–520, 2000.
- [26] L.R. Ford and D.R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8:399–404, 1956.
- [27] Bernard Fortz, Jennifer Rexford, and Mikkel Thorup. Traffic engineering with traditional ip routing protocols. *IEEE Communications Magazine*, pages 118–124, October 2002.
- [28] Bernard Fortz, Patrick Soriano, and Christelle Wynants. A tabu search algorithm for self-healing ring network design.
- [29] Bernard Fortz and Mikkel Thorup. Internet traffic engineering by optimizing OSPF weights. In *INFOCOM (2)*, pages 519–528, 2000.
- [30] A. Frank. Edge disjoint paths in planar graphs. *J. on Comb. Theory, Series B*, 38:164–178, 1985.
- [31] A. Frank. In: *Paths, Flows and VLSI Layout*, B. Korte, L. Lovász, H.J. Prömel, A. Schrijver (eds), chapter Packing paths, circuits and cuts - a survey. Springer Verlag, Berlin, 1990.
- [32] P.-J. Wan G. Calinescu. “Traffic partition in wdm/sonet rings to minimize sonet adms”. *Journal of Combinatorial Optimization*, 6:425–453, 2002.
- [33] H. N. Gabow, M. X. Goemans, and D. P. Williamson. An efficient approximation algorithm for the survivable network design problem. In “*3rd MPS IPCO conference*”, pages 57–74.
- [34] N. Garg and J. Konemann. “faster and simpler algorithms for multicommodity flow and other fractional packing problems”. “*Symposium on Foundations of Computer Science*”, pages “300–309”, 1998.
- [35] O. Goldschmidt, A. Laugier, and E.V. Olinick. Sonet/sdh ring assignment with capacity constraints. *Discrete Applied Mathematics*, 129(1):99–128, 2003.
- [36] R.E. Gomory and T.C. Hu. Multi-terminal network flows. *SIAM Journal of Applied Mathematics*, 9:551–570, 1961.
- [37] R.E. Gomory and T.C. Hu. Synthesis of a communication network. *SIAM Journal of Applied Mathematics*, 12:348–369, 1961.
- [38] M. Grötschel, C. Monma, and M. Stoer. Design of survivable networks. 7, 1995.
- [39] R. Hassin and A. Levin. Approximation algorithms for constructing wavelength routing problems. *Networks*, 40(1):32–37, 2002.

- [40] C.P.M. van Hoesel, A.M.C.A. Koster, R.L.M.J. van de Leensel, and M.W.P. Savelsbergh. Polyhedral results for the edge capacity polytope. *Mathematical Programming*, 92(2):335–358, 2002.
- [41] T.C. Hu. Multicommodity network flows. *Operations Research*, 11:344–360, 1963.
- [42] T.C. Hu. Optimum communication spanning trees. *SIAM Journal on Computing*, 3:188–195, 1974.
- [43] D. S. Johnson, J. K. Lenstra, and A. H. G. Rinnooy Kan. The complexity of the network design problem. *Networks*, 8:279–285, 1978.
- [44] D. Karp. On the complexity of combinatorial problems. *Networks*, 5:45–68, 1975.
- [45] J. Kennington, E. Olinick, A. Ortynski, and G. Spiride. Wavelength routing and assignment in a survivable wdm mesh network. *Operations Research*, 51(1):67–79, 2003.
- [46] S. Khanna. A polynomial time approximation scheme for the sonet ring loading problem. *Bell Labs Tech. J.*, 2, 1997.
- [47] J.M. Kleinberg and A. Kumar. Wavelength conversion in optical networks. *Journal of Algorithms*, 38(1):25–50, 2001.
- [48] A.M.C.A. Koster and A. Zymolka. Minimum converter wavelength assignment in all-optical networks. Technical Report ZIB-Report 03-45, Konrad-Zuse-Zentrum für Informationstechnik Berlin, 2003.
- [49] A.M.C.A. Koster, A. Zymolka, M. Jäger, R. Hülsermann, and C. Gerlach. Demand-wise shared protection for meshed optical networks. Preprint ZIB Report 03-24, Konrad-Zuse-Zentrum für Informationstechnik Berlin, 2003.
- [50] M.R. Kramer and J. van Leeuwen. Wire routing is np-complete. Technical Report "RUU-CS-82-04", Utrecht University, 1982.
- [51] M. Laguna. Clustering for the design of sonet rings in interoffice telecommunications. *Management Science*, 40(6):1533–1541, 1994.
- [52] R.L.M.J. van de Leensel. *Models and Algorithms for Telecommunication Network Design*. PhD thesis, Maastricht University, 1999.
- [53] T. Leighton, F. Makedon, Serge Plotkin, C. Stein, E. Tardos, and S. Tragoudas. Fast approximation algorithms for multicommodity flow problems. *Journal of Computer and System Sciences*, 50(2):228–243, 1995.
- [54] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of ethernet traffic. *IEEE/ACM Transactions on Networking*, 2(1):1–15, 1993.
- [55] A. van der Kraaij J. Lutton M. Bouhtou, S. van Hoesel. Linear tarification in multi-commodity telecommunications networks. Technical Report METEOR Research Memorandum 02-012, University Maastricht, 2002.
- [56] T.L. Magnanti, P. Mirchandani, and R. Vachani. "The convex hull of two core capacitated network design problems". *Mathematical Programming*, 60(2):233–250, 1993.

- [57] T.L. Magnanti, P. Mirchandani, and R. Vachani. “Modelling and solving the two-facility capacitated network loading problem”. *Operations Research*, 43(1):142–157, 1995.
- [58] T.L. Magnanti and S. Raghavan. Strong formulations for network design problems with connectivity requirements. Technical Report TR1999-28, Institute for systems research, 1999.
- [59] A.R. Mahjoub. Two edge connected spanning subgraphs and polyhedra. *Mathematical Programming*, 64:199–208, 1994.
- [60] K. Menger. Zur allgemeinen kurventheorie. *Fundamenta Mathematicae*, 10:96–115, 1927.
- [61] M. Minoux. “Network design and optimum network design problems: Models, solution methods and applications”. *Networks*, 19:313–360, 1989.
- [62] H. Okamura and P. Seymour. “Multicommodity flows in planar graphs”. *Journal of Combinatorial Theory B*, 31:75–81, 2002.
- [63] A.E. Ozdaglar and D.P. Bertsekas. Optimal solution of integer multicommodity flow problems with application in optical networks. Santorini, greece, Proc. of Symposium on Global Optimization, 2003.
- [64] A. Schrijver, P. Seymour, and P. Winkler. “The ring loading problem”. *SIAM J. Discrete Mathematics*, 11(1):1–14, 1998.
- [65] P.D. Seymour. In: *Graph Theory Related Topics*, J.A. Bondy and U.S.R. Murty (eds), chapter Sum of circuits, pages 341–355. Academic Press New York, 1990.
- [66] H.D. Sherali, J. C. Smith, and Y. Lee. Enhanced model representations for an intra-ring synchronous optical network design problem allowing demand splitting. *INFORMS Journal on Computing*, 12(4):284–298, 2000.
- [67] A. Sutter, F. Vanderbeck, and L. Wolsey. “Optimal placement of add/drop multiplexers: heuristic and exact algorithms”. *Operations Research*, 46(5):719–728, 1998.
- [68] Hoesel S. van and R. Müller. Optimization issues in combinatorial auctions. Technical Report METEOR Research Memorandum, University Maastricht, 2004.
- [69] Hoesel S. van, Kraaij A.F. van der, Mannino C., Bouhtou M., and Oriolo G. Polynomial cases of the tariffication problem. Technical Report METEOR Research Memorandum 03-063, University Maastricht, 2003.
- [70] G.T. Wilfong and P. Winkler. “Ring routing and wavelength translation”. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 98)*, pages “333–341”, 1998.
- [71] R. Wong. Worst-case analysis of network design problem heuristics. *SIAM Journal on Algebraic and Discrete Methods*, 1:51–63, 1980.
- [72] B.Y. Wu, G. Lancia, V. Bafna, K.M. Chao, R. Ravi, and C.Y. Tang. A polynomial-time approximation scheme for the minimum routing cost spanning trees. *SIAM Journal on Computing*, 29:761–778, 1999.

- [73] H. Zang, J. Jue, and B. Mukherjee. A review of routing and wavelength assignment approaches for wavelength-routed optical wdm networks. *Optical Networks Magazine*, 1(1):47–60, 2000.